DOI: 10.24996/ijs.2019.60.2.18

ISSN: 0067-2904

Arabic Handwriting Word Recognition Based on Scale Invariant Feature Transform and Support Vector Machine

Alia Karim Abdul Hassan*, Bashar Saadoon Mahdi, Asmaa Abdullah Mohammed
Computer Science Department, University of Technology, Baghdad, Iraq

Abstract
Offline Arabic handwritten recognition lies in a major field of challenge due to the changing styles of writing from one individual to another. It is difficult to recognize the Arabic handwritten because of the same appearance of the different characters. In this paper a proposed method for Offline Arabic handwritten recognition. The proposed method for recognition hand-written Arabic word without segmentation to sub letters based on feature extraction scale invariant feature transform (SIFT) and support vector machines (SVMs) to enhance the recognition accuracy. The proposed method experimented using (AHDB) database. The experiment result show (99.08) recognition rate.

Keywords: AHDB Database, SIFT feature extraction, SVM classifier algorithm.

التعرف على الكلمة العربية المكتوبة بخط اليد مرتكز على سمة ثابتة التغير والة متجه الدعم تحويل

علياء كريم عبد الحسن*، بشار سعدون مهدي، اسماء عبد الله محمد
كلية العلوم،الجامعة التكنولوجية، بغداد، العراق

الخلاصة
تعد أهمية التعرّف على الكتابة العربية بخط اليد من التحديات بسبب أنماط الكتابة المتغيرة من فرد إلى آخر. من الصعب التعرف على اللغة العربية المكتوبة بخط اليد بسبب تشابه مظهر الحروف المختلفة. في هذه الورقة، طريقة مقترحة للتعرف على الكتابة العربية بخط اليد بدون اتصال. الطريقة المقترحة للتعرف على الكلمة العربية المكتوبة بخط اليد دون التقسيم إلى أحرف فرعيه بناء على تحويل سمة ثابتة لاستخراج سمة تحويل ثابته(SIFT) وآلات متجه ال دعم(SVM) لتعزيز دقة التعرف ,جربت الطريقة المقترحة بأستخدام قاعدة البيانات،(AHDB) وكانت نسبة التعرف (99,08).

1. Introduction

Handwriting recognition is the process of converting the handwriting text images into a text file that understandable by the computer and used for many purposes[1]. To recognize handwritten words or characters there are several strategies in the computational pattern recognition such as artificial neural networks and statistical approaches like K-Nearest Neighbor KNN. Naturally, handwriting is cursive due to several factors which are the writer’s style, quality of paper and geometric factors controlled by the writing condition is very unsteady in shape and quality of tracing [2]. Optical Character Recognition (OCR), refers to the branch of computer science that involves reading text from paper and translating the images into a full editable form that the computer can manipulate. An OCR system enables you to take a hard copy of a book or a magazine article and scan it into an electronic computer file, and then edit the file using word processor applications [3].

*Email: 110018@uotechnology.edu.iq

The recognition system has several stages such as Image acquisition, preprocessing, feature extraction and classification / recognition. Achieving high recognition accuracy depends mainly on the used feature selection method [4]. The first stage in any recognition system is preprocessing which tries to reduce the noise data and keep only the desired information and make the next operation (feature extraction process) easy to implement. Moreover, the second stage is feature extraction and selection; they are extracted useful information from the binary handwriting, word image to be used in recognition stage. The last stage is classification and recognition, which classified all data in different classes, then recognize the unknown handwritten word image to desired class [5]. In this paper, a proposed algorithm for Arabic word handwritten recognition without segmented the word to sub latters in high accuracy using SIFT and SVM.

2. Related work

In 2015 [1] The proposed work depends on the handwriting word level, and it does not need for character segmentation stage. An Arabic handwriting dataset AHDB, dataset used for train and test the proposed system. Besides, the system achieved the best recognition accuracy 96.317% based on several feature extraction methods and SVM classifier. Experimental results show that the polynomial kernel of SVM is convergent and more accurate for recognition than other SVM kernels. In [5] the proposed method uses a 4 levels discrete wavelet transform (DWT) on binary images. Sliding window on wavelet space computes the stander derivation for each window. The extracted features were classified with multiple Support Vector Machine (SVM) classifiers. The experimental results of the simulation show 94.44% recognition rate. In [6] the research use 4-level of the Discrete Wavelet Transform (DCT) to extract the features from the word image, then fed into the K-nearest neighbor classifier (KNN). The recognition rate of this system was 50.83%. In [7] design a recognition system for Arabic words. The feature extraction stage used a Freeman code method which determined by the contour of the image, Zernike moments, and structural features.

The classification phase uses multiple classifiers such as K-Means algorithm, Probabilistic Neural Network, Fuzzy C-Means algorithm and K Nearest Neighbor algorithm. Also introduced a handwriting database of Algerian city names. The recognition rate was 80%. In [4] A proposed system to recognize Arabic word based on using DCT for feature extraction stage and support vector machine classifier for recognition stage. Recognition rate was 91.70% of the IFN/ENIT Arabic standard database. In [8] use DCT features which extracted from each word sample, then features are fed to train a neural network for classification. The proposed system tested on IFN/ENIT Database each time 80% of the samples in the database is used for training and the remaining 20% for testing recognition rate 82.5%. In [5] a proposed system for identifying handwritten word by using more than one technique for feature extraction of the image and used SVM classifier. To estimate the performance, own dataset is created and sample words for testing collected from multiple people.This presented work gives 88.13% of recognition rate.

3. Scale Invariant Feature Transform (SIFT) [9] [10].

Is normally utilized for feature extraction from images which have stable behavior in image rotation, image translation, image illumination, image scaling and different camera viewpoint, The algorithm main computation steps are:

Scale space construction: taking the original image and generating a sequence of continuously blurred images, then resizing the original image to 50% of its size and generating blurred-out images again and again. The scale space of an image is represented in a form of a function of L(x, y, σ), which is generated from convolving a variable-scale Gauss, G(x, y, σ), with an input image, I(x, y): L(x, y, σ) = G(x, y, σ) ∗ I(x, y), using equation-1.

$$\mathbf{G(x, y, \sigma) = \frac{1}{2\pi\sigma 2}\, e^{-(x2+y2)/2\sigma 2}} \tag{1}$$

Where L denotes a blurred image, G denotes the Gaussian blur operator, I denotes the image, x, y denote coordinates of the index, σ denotes the parameter of the “scale” (the degree of blurring). The bigger the value, the greater is the blurring, the * denotes the convolution process in x and y.

Difference Of Gaussian(DOG) Approximation: The scale space is used to calculate the difference between two consecutive scales (Difference of Gaussian) the DoG function convolved with the image, D(x, y, σ), which may be calculated from subtracting two neighboring scales that are separated by a constant multiplicative factor k using equation-2. There are a many reasons to choose this formula. It's a specifically sufficient formula to calculate, due to the fact that the smoothed images, L, must be calculated in any case, for a scale space feature description, and D may thus be calculated by simply subtracting images.

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (2)$$

Find key points: locate maxima/minima in DOG image and find sub pixel maxima/minima. X is a key point if its value is the biggest or smallest of all 26 neighbors in 3x3 regions. This to find the "Maxima and minima" of the Gaussian images.
Get rid of bad key points: Edges and low contrast areas are considered as bad key-points. The elimination of those key points increases the algorithm's efficiency and robustness. An approach which is similar to the Harris Corner Detector has been utilized here.
Assigning orientations of the key points: An orientation is obtained for every key point. Additional computations are performed based this orientation. Which effectively rid of the orientation impact, making it invariant to rotations. Gradient magnitudes and orientations are obtained with the use equations3&4. The magnitudes and orientations are calculated for all of the pixels that surround the key points. After that, a histogram is generated.

$$m(x,y)=\sqrt{L(x+1,y)-L(x-1,y))2+(L(x,y+1)-L(x,y-1))2} \quad \textbf{(3)}$$

$$\theta(x,y)=\tan^{-1}\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \quad \textbf{(4)}$$

Produce SIFT features: lastly, having the scale and rotation invariance, one more representation is produced. Which is helpful in the unique identifications of features. With this representation, it is possible easily identifying the required features (for example, a certain eye, or a sign board). To do this, a 16×16 window that surrounds the key point is set and this window is divided into 16 windows of size 4x4,shown in Figure-1. Inside every 4x4 window, gradient magnitudes and orientations are obtained. This histogram is split to eight bins and the amount of orientation that is added to the bin is dependent on the gradient magnitude. Finally, every key point is defined by 4*4*8= 128 dimensional feature vector.

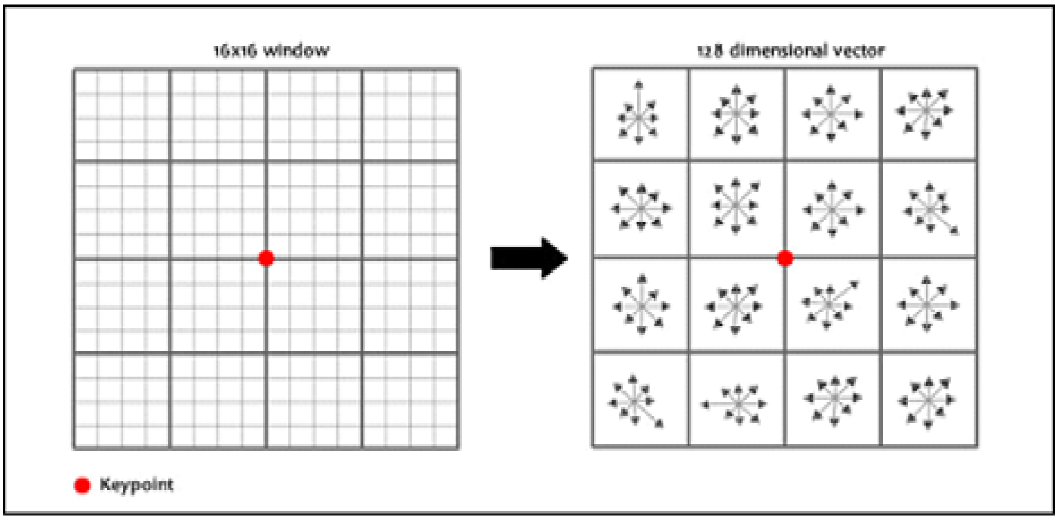

Figure 1-SIFT Features [11]

4. Bag of Word (BOW)

One of the most generalized and commonly utilized approaches for class recognition is the bag of words (BOW) also referred to as the bag of features or bag of key point model. This approach produces a histogram, which is distributing visual words that are obtained from the test image, and then classifiers perform a classification of the image according to every classifier's of features **[12]**.

5. K-means clustering [13]

K-means clustering an approach of clustering observations to a certain number of distinct clusters "K" represents the number of determined clusters specified. There are different distance measurements to determine what observation to append to what cluster. The approach has the aim of minimizing the measurement between the cluster's centroid and the specific observation via the iterative appending of an observation to any of the clusters and terminating when achieving the minimal distance measurement (the algorithm-1 list the main steps).

Algorithm 1:Implement k-means **Input**: key points of image, number of the clusters. **Output**: Key point with cluster number length.
Begin Step1: Number of clusters k Step2: Find random centroid for key points. Step3: Calculate distance between key points and centroid by Equation – 5 Step4: Grouping key points according to minimum distance. Step5: Repeat steps 2,3 and 4 until the centroids no longer move. **End**

6. Fast Approximate Nearest Neighbor FANN[14]

FANN is utilized for performing the matching procedure on the clusters which have resulted from the preceding process, the advantage is that the use of the clusters will be faster and more precise in matching than separate features.

7. Support Vector Machine (SVM) [15]

It is an arrangement of related directed learning approaches that are used for order and forecast. Simply speaking, considering an arrangement of preparing instances, every set separate as having a place with one of two classifications. SVM produces a hyperplane arrangement of hyperplanes in a high or interminable dimensional space that may be used to characterize various tasks. Instinctively, a considerable deal is performed by the hyperplane which has the biggest separation for the nearest preparing data purposes of any of the classes (supposed useful edge), due to the fact that the bigger the edge the lower is the classifier's speculation mistake.

8. Proposed Arabic Handwriting Word Recognition System

Figure-2 shows the general view of the proposed system for Arabic handwritten word recognition. The presented approach for handwritten word identification without segmented to sub letters has several major steps are, preprocessing, feature extraction bag of word (K-mean, fast approximate Nearest neighbor), classification and recognition, and finaly Output :decision. Algorithm-1 list the detailed description of the proposed main steps.

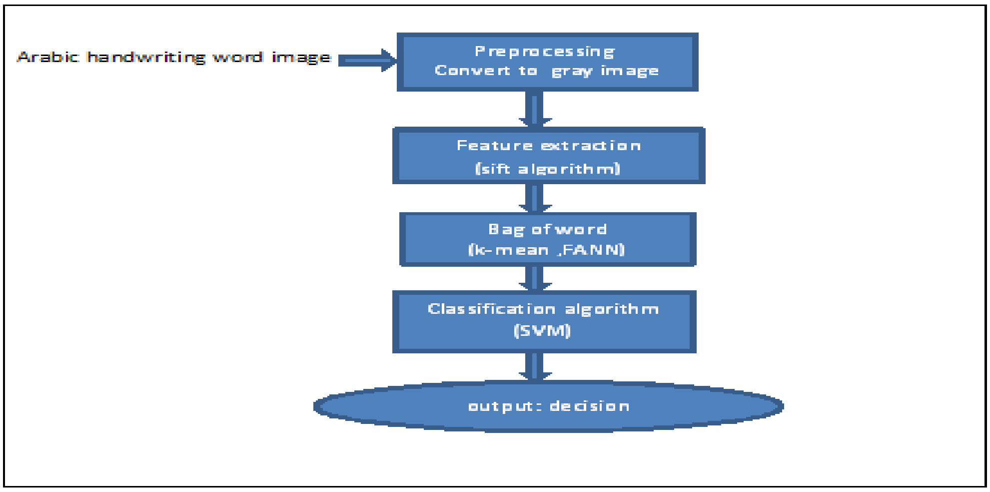

Figure -2 :Proposed system general steps

In **a**lgorithm-1 *step,1* preprocessing image most of the identifying and classification techniques require the data to be in a predefined type in this model we convert e color image (RGB) to gray image to easy implementation in the next step. In *step2*, **"**feature extraction is a very necessary step in word identification systems and for a wide range of pattern identification tasks. It has the aim of deleting the redundancy of data and obtain a better representation of the text image. *Step3*, "cluster extracted features into groups using k-means. Features divided into k clusters where every one of the features is part of the cluster of its closest average. The clustered properties are to be used for generating the histogram. In this process, every patch with an image is assigned to a specific code word via the procedure of k-means clustering and therefore, every one of the images can be represented with a histogram of the code words. In the *fourth step*, for the purposes of matching fast approximate Nearest Neighbors (FANN) algorithm has been utilized in order to obtain those property clusters have to be matched with each feature cluster in the database. This is the ultimate step prior to performing the classification itself. Finaly at *step five* nonlinear **SVMs** have been used for the inputs to classify them into accept or not. Figure-3 show example for implementing the steps of the proposed system

Algorithm-1: Proposed Arabic handwriting, System
Input: word image**s.**
Output: the Arabic handwriting, word image class label.

Step1: Preprocessing input image and convert color image(RGB) to gray image.
Step2: To delete the redundancy of data and obtain a better representation of the text image, apply SIFT algorithm on Arabic handwriting word images to extract feature .
Step3: cluster extracted features into groups using k-means. Use clustered properties for generating
the histogram. Every cluster of the images can be represented with a histogram of the code words.
Step 4:Apply fast approximates FANN to classify the features in the cluster without repeated.
Step5: Apply support vector machine (SVM) to classifier the Arabic handwriting word image .
Step6: return(class label)

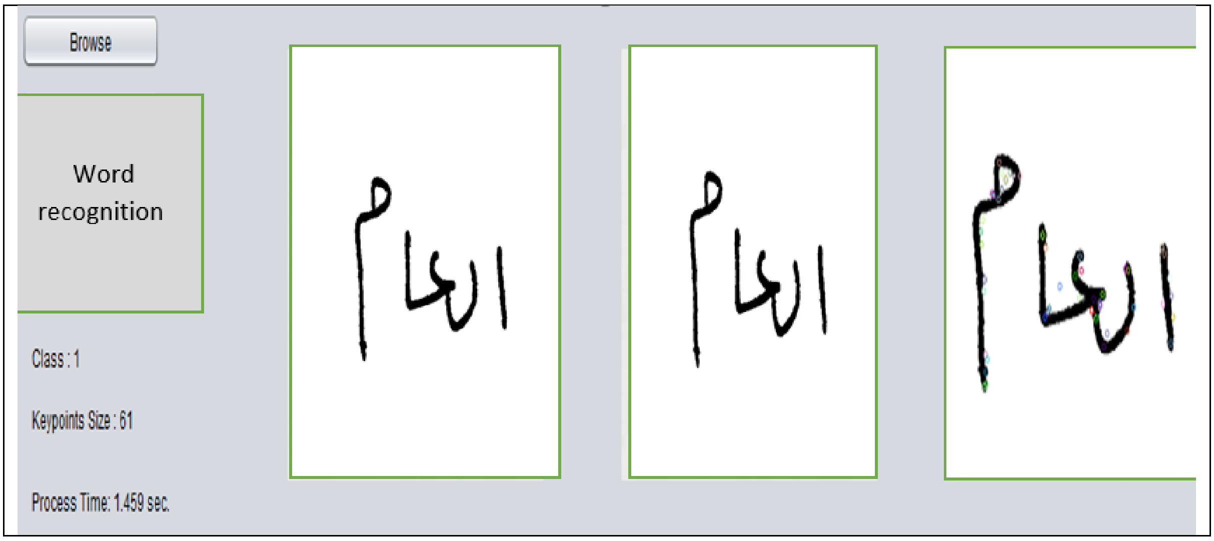

Figure 3-Example for implementing the steps of the proposed system

9. Experimental Result

The suggested approach has been implemented with the use of the visual c++ R2013a, under windows-7 64-bit OS, with RAM 4GB, CPU 2.50GHz core i5 and it proved to be less time consuming and more efficient. This work is experimented on a dataset (AHDB) that includes 4700 handwritten images which contain the most common Arabic words written by hundreds of different writers [16]. 70 % were used for the training data and 30 %for testing. The results were evaluated using the traditional evaluation measure, Precision (P) is the fraction of retrieved documents that are relevant (equation-6).

Precision=$\frac{\text{(relevant items retrieved)}}{\text{retrieved items}}$ ....(6)

[17] Table -1 shows the comparative results of the proposed system with the most related works using AHDB database with 2072 for training set and 868 for testing set. From Table-1 it noted that the highest accuracy obtained by the proposed system. The processing time of the proposed system was *7.266 msec.*

Table1-shows the comparative results of the proposed system with the most related works.

Related work	Years	Techniques	Accuracy
[8]	**2015**	**SVM+densityfeatures, long run features and structural features**	**88.31%**
[1]	**2015**	**SVM**	**96.317%**
[2]	**2017**	**Hybrid (SVM, KNN)**	**94.29%**
The proposed system	**2018**	**SVM+SIFT**	**99.08%**

10. Conclusion

Handwritten recognition is a challenging problem for researchers because the text was written by various writers and has a high level of ambiguity and complexity. It plays essential roles in many applications, such as office automation, checks, verification, mail sorting, and a large variety of banking, business as well as natural human-computer interaction. The proposed system, identifying the handwritten Arabic word as one entity without segmentation to sub letters, Feature extraction is a very significant stage in word identification systems and for a great deal of pattern identification tasks. It has the aim of removing the redundant objects from data and obtain a more efficient representation of text images of a group of numerical characteristics with the use of (SIFT) algorithm, feature selection used (K-Mean, Fast Approximate Nearest neighbor FANN). In this paper, we used (SVM) which is one of the soft computing techniques for classification. It has been noticed during the process that the rate of success of any identification system is not merely dependent on the step of feature extraction, but it is also dependent on numerous reasons like the preprocessing step and the classifier, recognizer technique. Implementation with the dataset (AHDB) obtained precision was 99.08% and processing time7.266 m sec. The future Work enhancing the presented system to be capable of working in real time via applying it with on-line recognition of Mobile and tablet devices.

References

1. Kadhm, Mustafa S., and Asst Prof Dr Alia Karim Abdul. 2015. Handwriting Word Recognition Based on SVM Classifier. (IJACSA) *International Journal of Advanced Computer Science and Applications*, **6**(11): 64-68.
2. Abdul Hassan, A. K., Kadhm. M.S. **2016**. Arabic Handwriting Text Recognition Based on Efficient Segmentation, DCT and HOG Features. *International Journal of Multimedia and Ubiquitous Engineering*, **11**(10): 83-92. http://www.earticle.net/article.aspx?sn=288474.
3. Al-Garalleh, B. A. A. W. **2013**. A Framework Model for Arabic Handwritten Text Recognition to Handle Missing Text Fragments (Doctoral dissertation, Faculty of Information Technology Middle East University Amman – Jordan). pp: 127.
4. Hirwani, A., Verma, N. and Gonnade, S. **2014**. Efficient Handwritten Alphabet Recognition Using LBP based Feature Extraction and Nearest Neighbor Classifier. *Int. J. Adv. Res. Comput. Sci. Softw. Eng*., **4**(11): 2277–128, pp.549-553.
5. Abdul Hassan, A.K. and Alawi, M. **2017**. Proposed Handwriting Arabic Words classification Based On Discrete Wavelet Transform and Support Vector Machine, *Iraqi Journal of Science*, **58**(2C): 5511 -1168,DOI: 10.24996.ijs.58.2C.19.
6. AlKhateeb, J. H. **2010**. Word Based Off-line Handwritten Arabic Classification and Recognition. MSc thesis. University of Bradford Open Access repository.
7. Nemouchi1, S., Souici M. L. and Farah, N. **2012**. Classifiers Combination for Arabic Words Recognition Application to Handwritten Algerian City Names. International Conference on Image and Signal Processing. Springer, Berlin, Heidelberg.7340:562-570**.**
8. AlKhateeb, J. H. **2011**. Word-Based Handwritten Arabic Scripts Recognition Using Dynamic Bayesian Network. The 5th International Conference on Information Technology.pp:6.

9. Shurth, A. and Patel, M. S. **2015**. Offline Handwritten Word Recognition using Multiple Features with SVM Classifier for Holistic Approach, **3**(6): 5989–5995.

10. Lowe, D.G. **2004**. Distinctive image features from scale invariant keypoints. International Journal of Computer Vision Manufactured in The Netherland. **60**(2): 91–110.

11. https://image.slidesharecdn.com/sift1-120328115134-phpapp02/95/sift-21-728.jpg?cb=1347276683.

12. Kim, J., Kim, B.S. and Savarese, S. **2012**. Comparing Image Classification Methods: K-Nearest-Neighbor and Support-Vector-Machines. *Appl. Math. Electr. Comput. Eng*. **1001**: 48109-2122

13. Yi, Y. and Dong, x .**2010**. Image Clustering Using Local Discriminant Models And Global Integration.*IEEE transaction on image processing*. **1**(10):2761-2773.

14. Marius Muja and David G. Lowe. **2014**. Scalable Nearest Neighber Algorithms for High Dimensional Data. in *IEEE Transactions on Pattern Analysis & Machine Intelligence*, **36**(11): 2227-2240. doi:10.1109/TPAMI.2014.2321376.

15. Aliev, R. A. and Aliev, R.R. **2001**. Soft computing and its applications. *Word scientific Singapore. New jersey*. *London. Hong Kong*. **144**(1): 365–366.

16. Al-Ma’adeed S. Al-, Elliman, D. and Higgins C. **2004**. A Data Base for Arabic Handwritten Text Recognition Research. *The International Arab Journal of Information Technology*, **1**(1): 117-121.

17. Manning, C.D., Raghavan, P. and Schütze, H. **2009**. Chapter 8: Evaluation in information retrieval. *Introduction to Information Retrieval* .: 1–18.